

MEDICAL RESEARCH DEPARTMENT



U. S. Submarine Base New London

AN INVESTIGATION OF THE BILLET QUALIFICATIONS BLANK
(Navpers 16418) FOR ITS SUITABILITY WITH COMBAT-
EXPERIENCED CANDIDATES FOR SUBMARINE TRAINING.

By

Lt. (jg) N. R. Bartlett, H(S), USNR

First and final Report

Bureau of Medicine and Surgery

Research Project No. X-598 (Sub. No. 129)

14 January 1946

APPROVED: Captain C. W. Shilling, (MC) U. S. Navy, MO-in-C

APPROVED FOR PUBLIC
RELEASE - DISTRIBUTION
UNLIMITED

LIBRARY

MEDICAL RESEARCH LABORATORY

ACCESSION NO. FILE COPY # 2

AN INVESTIGATION OF THE BILLET QUALIFICATIONS BLANK
(Navpers 16418) FOR ITS SUITABILITY WITH COMBAT-
EXPERIENCED CANDIDATES FOR SUBMARINE TRAINING.

By

Lt. (jg) N. R. Bartlett, H(S), USNR

Final Report

Bureau of Medicine and Surgery

Research Project No. X-598 (Sub. No. 129)

"Investigation of the BQB"

Medical Research Department

U. S. Submarine Base

New London, Connecticut

14 January 1946

SUMMARY AND CONCLUSIONS

1. Distributions of scores for a combat-experienced population of candidates for submarine duty are presented for the several sections of the Billet Qualifications Blank (Navpers 16418) and of the two sections of the Enlisted Personal Inventory (Navpers 16845).

2. Scattergrams for sections of the Billet Qualifications Blank with Part I of the Personal Inventory and with the Navy General Classification test are tabulated for this population.

3. Data on the success of the two personality tests in predicting certain criterion examination classifications of combat-experienced submarine candidates are presented. The tables are intended as supplements to similar research on recruit populations.

4. It is evident that both the Personal Inventory and the Billet Qualifications Blank tap something relevant in the criterion classifications. However, the population reported herein is too small to warrant any extensive conclusions on the relative significance of the two instruments.

5. The combination of the two personality tests certainly is related to failure in examinations for submarine training. This fact warrants the assumption that personality tests can be utilized to good effect with combat-experienced populations.

INVESTIGATION OF THE BILLET QUALIFICATIONS BLANK

BACKGROUND

The Billet Qualifications Blank is a "personality" test designed to identify neurotic recruits. It has one particular feature differentiating it from typical tests involving forced-choice responses and check-lists of psychosomatic complaints; that feature is the following: an individual is asked to describe himself by choosing from a long list of job or billet characteristics those which describe jobs he believes himself suited to perform. The job descriptions are couched in terms that make his choices depend upon personality factors.

This activity has collaborated with the Bureau of Naval Personnel in a study of the validity of the Billet Qualifications Blank for screening recruits for assignment to submarine duty. Recruits at training centers were examined with both the Billet Qualifications Blank and the Navy Enlisted Personal Inventory. The answer sheets were forwarded to the Bureau of Naval Personnel, and the Submarine Base informed the Bureau of the success of each in meeting Medical Research Department requirements for acceptance and school requirements for graduation. This information served to evaluate the two tests. In order to avoid prejudice in the criterion, Medical Research examinations were conducted entirely independently of psychological personality tests; the decision to accept or reject the candidate was a clinical judgment exercised without the aid of any scores except for tests of the Navy Basic Battery for training aptitude.

This procedure was set to furnish a criterion for judging the screening effectiveness of tests with recruits; and the data on recruits are now under analysis by BuPers. However, men who have experienced combat duty certainly present a different attitude toward what they term "nut" tests. Furthermore, the basis on which experienced personnel are judged may be markedly different from the basis for judging the probable military effectiveness of recruits. For this reason, an instrument developed and proved on recruit populations may be a complete waste of time with combat-experienced personnel. Fortunately, though, one can take advantage of the same procedure described above to judge the screening effectiveness with a population comprised of men of the latter type. Nothing is added to the experiment but the time and labor involved in a half hour of additional examination of each of the candidates reporting for submarine training from surface craft.

This paper extends the research on the Billet Qualifications Blank by adding data on the effectiveness of the test in screening combat-experienced candidates for submarine duty. Score distributions for the various parts of the test are presented, and the relationships of the scores to certain criterion classifications are tabulated.

METHOD

A system was established on May 1, 1945, whereby no test scores other than for the Navy Basic Battery were made available to the officers conducting examinations for Submarine School. Each man for whom the service record indicated previous sea duty was required to complete two personality schedules -- the Billet Qualifications Blank and the Navy Enlisted Personal Inventory -- upon reporting for examination. Order of testing was alternated day-by-day. Examination results then were grouped into four categories according to the psychiatric and medical action taken in each instance. These categories were: I, Accepted (passed all psychometric, medical, psychiatric and training standards); II, Psychiatric Failures; and IV, Physical, Psychometric, and Training Failures. The three failing categories are more or less self-explanatory. If a man failed the physical or did not meet the training standards he was classified in group IV. If he requested to be disqualified because he no longer desired submarine duty and such a change of intention was not regarded by the psychiatrist as symptomatic of a fundamental instability, he was considered a simple motivation. (Group III). And finally, if his general personality or emotional habit structure was judged markedly inadequate he was classified a psychiatric failure. (Group II).

RESULTS

Scattergrams of each of the four classifications with the scores on the several parts of the two tests are presented in Table I on the next page. Unfortunately, the numbers in the four categories are too small to allow much insight into relationships. But it is emphasized that every one of the sections appears to tap something relevant in the examination classifications.

TABLE I

Scattergram showing the relationship of section scores
to Examination Classifications

Score Range BQB Part I	Examination Classifications				Total
	I	II	III	IV	
0-1			1		1
2-3					0
4-5					0
6-7					0
8-9	1			1	2
10-11					0
12-13				1	1
14-15					0
16-17	2	1	1		4
18-19	1	2			3
20-21	4			1	5
22-23	7	2	1	1	11
24-25	3	4	1		8
26-27	9			5	14
28-29	11	2		4	17
30-31	7	1		3	11
32-33	9		2	3	14
34-35	23	1	2	5	31
36-37	27	3		6	36
38-39	29		1	3	33
40-	4	1		2	7
Total	137	17	9	35	198
BQB Part II					
8-9			1		1
10-11		1		1	2
12-13	2	2	2	1	7
14-15	5	2	1	1	9
16-17	11	3	1	5	20
18-19	38	3	2	10	53
20-21	43	4		7	54
22-23	37	1	2	9	49
24-25	1	1		1	3
Total	137	17	9	35	198

TABLE I (continued)

BQB Part III	I	II	III	IV	Total
8-9					0
Poor 10-11				1	1
12-13		1			1
14-15		2	2		4
16-17	7	4	1	7	19
18-19	36	6	1	8	51
Good 20-21	49	2	3	12	66
22-23	40	1	1	5	47
24-25	5	1	1	4	11
Total	<u>137</u>	<u>17</u>	<u>9</u>	<u>35</u>	<u>198</u>

BQB Part IV	I	II	III	IV	Total
0	14	3		6	23
1	27	1	1	4	33
2	17	4		4	25
3	17		3	3	23
4	16	5	2	8	31
5	11	1	1	3	16
6	10		1	3	14
7	10	2		2	14
8	11				11
9	2			2	4
10	2	1	1		4
Total	<u>137</u>	<u>17</u>	<u>9</u>	<u>35</u>	<u>198</u>

PI Part I	I	II	III	IV	Total
12-13			1		1
Poor 10-11			1	1	2
8-9	1		1		2
6-7	6	4		5	16
4-5	11	3	1	3	18
Good 2-3	41	6	2	12	61
0-1	78	4	3	15	93
Total	<u>137</u>	<u>17</u>	<u>9</u>	<u>35</u>	<u>198</u>

PI Part II	I	II	III	IV	Total
5			1		1
4					0
Poor 3				2	2
2	1	3	1	2	7
1	17	1	1	2	21
Good 0	119	13	6	29	167
Total	<u>137</u>	<u>17</u>	<u>9</u>	<u>35</u>	<u>198</u>

The relative screening efficiencies of the Personal Inventory and of the Billet Qualifications Blank can not be determined on the basis of the small amount of data available in this study; these data, after all, are only supplementary. From them alone one can say neither that one test is more valuable than the other, nor that either or both is effective enough to warrant addition to a battery. However, the problem of the extent to which the two tests cover the same characteristics can be touched upon by examining the degree of interrelationship between scores on the two. Table II presents scattergrams of Part I scores of the Personal Inventory with the four section scores of the Billet Qualifications Blank. Although the correlation is considerable in the first three of the scattergrams it is obvious that the overlap is far from complete in any single one. In all probability this is attributable to the multiplicity of factors involved in each section or part, as well as to imperfect reliability.

TABLE II

Relationship between Part I of the Personal Inventory
and the Sections of the BQB.

BQB Section I	Part I Score Ranges							Total
	0-1	2-3	4-5	6-7	8-9	10-11	12-13	
0-5							1	1
6-11				2				2
12-17	1		1		1	2		5
18-23	4	5	5	5				19
24-29	15	16	5	3				39
30-35	28	22	4	4				58
36-41	50	18	3	2	1			74
Total	98	61	18	16	2	2	1	198
BQB Section II	Part I Score Ranges							Total
	0-1	2-3	4-5	6-7	8-9	10-11	12-13	
8-9							1	1
10-11				1		1		2
12-13		1	2	2	1	1		7
14-15	2	3	2	2				9
16-17	7	9	2	2				20
18-19	24	21	2	5	1			53
20-21	31	16	5	2				54
22-23	31	11	5	2				49
24-25	3							3
Total	98	61	18	16	2	2	1	198

TABLE II - Continued

BQB Section III	0-1	2-3	4-5	6-7	8-9	10-11	12-13	Total
8-9								0
10-11				1				1
12-13				1				1
14-15	1	1				1	1	4
16-17	4	6	5	3		1		19
18-19	23	17	5	5	1			51
20-21	34	22	6	3	1			66
22-23	28	12	2	3				45
24-25	8	3						11
Total	98	61	18	16	2	2	1	198

BQB Section IV								
0	13	5	3	2				23
1	16	10	3	3	1			33
2	15	3	4	3				25
3	12	10			1			23
4	15	11	2	3				31
5	7	5		2		2		16
6	7	6	1					14
7	7	4	2	1				14
8	4	4	2	1				11
9	1	3						4
10	1		1	1			1	4
Total	98	61	18	16	2	2	1	198

Several studies have shown that the Personal Inventory is little related to measures of verbal and mechanical ability; on the other hand, it is useful for differentiating from a total population a group which will contain a larger proportion of psychiatric failures than does the rest of the population.¹ By a combination of these two circumstances it thus adds a significant element to a general program of aptitudes tests. Perhaps the Billet Qualifications Blank also may differentiate psychiatric failures. If so, before the instrument can be adopted, the same question of overlap with other aptitude tests must be explored.

The relationship of the parts of the Billet Qualifications Blank to the General Classification Test for the combat-experienced population of this study can be discerned in the scattergrams following. (Table III).

¹ OSRD Report #3963, "Final Report in Summary of Research on the Personal Inventory and Other Tests", 1 August 1944.

TABLE III

Relationship of BQB Section Scores to GCT Scores

BQB Part I	GCT Score Ranges									Total
	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	
0-4					1					1
5-9				1			1			2
10-14		1								1
15-19				1		3	3			7
20-24			2	8	5	2	1			18
25-29	1	1	2	11	2	7	7	4	2	37
30-34	2	1	2	11	10	8	7	1	1	43
35-39	1	2	6	21	16	16	11	7	2	82
40-45				2	3	1	1			7
Total	<u>4</u>	<u>5</u>	<u>12</u>	<u>55</u>	<u>37</u>	<u>37</u>	<u>31</u>	<u>12</u>	<u>5</u>	198
BQB Part II										
6-8					1					1
9-11				1			1			2
12-15		1	1	6	2	3	1	1		15
16-19	2	3	2	21	16	16	8	3	2	73
20-22	2	1	9	21	13	14	17	7	5	87
23-25				6	5	4	4	1		20
Total	<u>4</u>	<u>5</u>	<u>12</u>	<u>55</u>	<u>37</u>	<u>37</u>	<u>31</u>	<u>12</u>	<u>5</u>	198
BQB Part III										
9-11			1		1					2
12-15			1	2	1		1			5
16-19	4	3	4	22	13	12	8	3	1	70
20-22		1	5	27	18	18	15	5	2	91
23-25		1	1	4	4	7	7	4	2	30
Total	<u>4</u>	<u>5</u>	<u>12</u>	<u>55</u>	<u>37</u>	<u>37</u>	<u>31</u>	<u>12</u>	<u>5</u>	198
BQB Part IV										
0-1	4	2	7	20	9	8	6			56
2-6		3	4	29	22	20	16	11	3	108
7-10			1	6	6	9	9	1	2	34
Total	<u>4</u>	<u>5</u>	<u>12</u>	<u>55</u>	<u>37</u>	<u>37</u>	<u>31</u>	<u>12</u>	<u>5</u>	198

One notable finding is the positive correlation between scores for Section IV and scores for the General Classification Test. Section IV yields one index of the attitude of the man completing the test; a zero score is made by one who makes ridiculous claims of his ability and a high score by one who claims little. A casual interpretation of the correlation is that very low intelligence characterizes the men who choose extravagant descriptions of their abilities, and high intelligence many who are modest in their choices. The relationship between scores for this section and GCT scores is significant, at any rate.

In the analysis to follow, possibilities for establishing combination test screens are canvassed. It would appear from Table I that the relative number of failures could be reduced at a low cost in "passes" by using either of the two personality schedules. Table IV, however, suggests that perhaps the GCT alone might accomplish most of the screening. This possibility should be considered in conjunction with the correlation between GCT and Qualifications Blank scores.

The use of measures of training aptitude, characterizes many military selection programs. In view of the possibility that in a given situation, one cannot be content to weigh the effectiveness of a different instrument by noting the score distribution of criterion categories which it yields in that situation. Rather the problem must be considered in the light of whether the new test improves the screening effectiveness either by adding appreciably to the conventional aptitudes test screening in combination with such tests, or by offering better screening alone. For a number of reasons, it may be presumed that selection on the basis of training aptitudes tests will continue in effect; so for all practical purposes, the question is narrowed to whether the new instrument adds to the conventional screening. Table V shows the consequences of cuts with personality test scores after application of a minimum GCT score.

TABLE V

SUCCESS OF VARIOUS TEST SCREENINGS AFTER APPLICATION OF GCT SCORE MINIMUM

Type of screen	Detail of screen	Examination Classification Distribution				
		I	II	III	IV	Total
GCT	50 and above; Pass	151	15	7	24	177
	49 and below; Fail	6	2	2	11	21
	Total	137	17	9	35	198
GCT Plus PI Part II	GCT 50 and above & PI II-0 or 1 - Pass	130	12	5	21	168
	GCT 40 or below, & PI 2 or above: Fail	7	5	4	14	30
	Total	137	17	9	35	198
GCT Plus PI Part I	GCT 50 and above, & PI-I, 0 to 6: Pass	128	12	5	22	167
	GCT 49 or below, &/or PI 7 or above: Fail	9	5	4	13	31
	Total	137	17	9	35	198
GCT Plus PI Part I and PI Part II	GCT 50 & above: PI Part I, 0 to 6 & PI Part II 0 to 1: Pass	127	11	5	20	163
	GCT 49 or below, &/or PI Part I, 7 or above, and/or Part II, 2 or above: Fail	10	6	4	15	35
	Total	137	17	9	35	198
GCT Plus BQB	GCT 50 and above: BQB-I, 20 or above; II, 14 or above, and III, 16 or above: Pass	128	9	5	21	163
	GCT 49 or below, BQB below limits: Fail	9	8	4	14	35
	Total	137	17	9	35	198
GCT Plus PI Part I & II plus BQB	GCT, PI and BQB all above limiting cuts: Pass	126	7	5	20	158
	GCT, PI and BQB below limiting cuts: Fail	11	10	4	15	40
	Total	137	17	9	35	198

For illustration, consider the 25 men who fail the combination of BQB and GCT; it may be seen from the total of the second column that 21 of these men fail the GCT; of the remaining 14 only 3 meet the submarine criterion. In other words, the combination screens out men who are likely to fail. One can readily make the same analysis for the PI test, and, in general, the data are comparable for the two screens.

It is especially noteworthy that a combination of all three tests is the most successful in this study. In the next to bottom line of Table V, it may be seen that, of the 40 men who fail all three tests, 29 were rejected. In order to delineate further the contribution of the personality tests to this battery, Table VI was constructed.

TABLE VI

Scattergram showing success of experienced candidates in meeting examination standards for submarine duty according to results of a psychological test battery.

	Psychiatric Failures	Motivation Failures	Physical & Aptitude Fails.	Pass Tot.	
Fail GCT minimum.	2	2	11	6	21
Fail Personality tests, but pass GCT standard.	8	2	4	5	19
Pass GCT and Personality test standards.	7	5	20	126	158
Total	17	9	35	137	198

The overall improvement in the population which would have accrued through the use of a battery comprised of the research tests in conjunction with the General Classification Test is indicated in Table VI. This table shows the personality tests to have some validity. But more important, it emphasizes the need for incorporating psychological test methods into the system for assigning combat-experienced candidates to submarine school. Of 198 such candidates, 61 failed the examination standards after being transferred to New London from various activities scattered over the world. Had some test system been in effect, the proportion of psychological failures certainly would have been markedly lower.

Table VI suggests that a battery can be devised which, at a total cost of from five to ten per cent of the successful combat-experienced candidates, might eliminate approximately half the psychiatric failures. In other words, it is entirely appropriate to believe that paper and pencil "nut" tests will work with combat-experienced populations with much the same success reported for recruit populations.

Data for recruit populations on the same tests have been submitted to the Bureau of Naval Personnel for analysis.